



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# A genome resequencing-based genetic map reveals the recombination landscape of an outbred parasitic nematode in the presence of polyploidy and polyandry

### Citation for published version:

Doyle, S, Laing, R, Bartley, D, Britton, C, Chaudhry, U, Gilleard, J, Holroyd, NE, Mable, B, Maitland, K, Morrison, A, Tait, A, Tracey, A, Berriman, M, Devaney, E, Cotton, J & Sargison, N 2018, 'A genome resequencing-based genetic map reveals the recombination landscape of an outbred parasitic nematode in the presence of polyploidy and polyandry', *Genome Biology and Evolution*, vol. 10, no. 2, pp. 396-409.  
<https://doi.org/10.1093/gbe/evx269>

### Digital Object Identifier (DOI):

[10.1093/gbe/evx269](https://doi.org/10.1093/gbe/evx269)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Genome Biology and Evolution

### Publisher Rights Statement:

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# A Genome Resequencing-Based Genetic Map Reveals the Recombination Landscape of an Outbred Parasitic Nematode in the Presence of Polyploidy and Polyandry

Stephen R. Doyle<sup>1,\*</sup>, Roz Laing<sup>2</sup>, David J. Bartley<sup>3</sup>, Collette Britton<sup>2</sup>, Umer Chaudhry<sup>4</sup>, John S. Gilleard<sup>5</sup>, Nancy Holroyd<sup>1</sup>, Barbara K. Mable<sup>2</sup>, Kirsty Maitland<sup>2</sup>, Alison A. Morrison<sup>3</sup>, Andy Tait<sup>2</sup>, Alan Tracey<sup>1</sup>, Matthew Berriman<sup>1</sup>, Eileen Devaney<sup>2</sup>, James A. Cotton<sup>1,\*†</sup>, and Neil D. Sargison<sup>4,\*†</sup>

<sup>1</sup>Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, United Kingdom

<sup>2</sup>Institute of Biodiversity Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, United Kingdom

<sup>3</sup>Moredun Research Institute, Pentlands Science Park, Penicuik, United Kingdom

<sup>4</sup>Royal (Dick) School of Veterinary Studies, University of Edinburgh, United Kingdom

<sup>5</sup>Department of Comparative Biology and Experimental Medicine, Faculty of Veterinary Medicine, University of Calgary, Alberta, Canada

<sup>†</sup>These authors contributed equally to this work.

\*Corresponding authors: E-mails: sd21@sanger.ac.uk; jc17@sanger.ac.uk; neil.sargison@ed.ac.uk.

Accepted: December 15, 2017

**Data deposition:** The raw sequencing data generated and/or analyzed during this study are available in the European Nucleotide Archive repository, <http://www.ebi.ac.uk/ena/> under the study accession number ERP024253. The genome assembly is available at [ftp://ngs.sanger.ac.uk/production/pathogens/Haemonchus\\_contortus](ftp://ngs.sanger.ac.uk/production/pathogens/Haemonchus_contortus).

## Abstract

The parasitic nematode *Haemonchus contortus* is an economically and clinically important pathogen of small ruminants, and a model system for understanding the mechanisms and evolution of traits such as anthelmintic resistance. Anthelmintic resistance is widespread and is a major threat to the sustainability of livestock agriculture globally; however, little is known about the genome architecture and parameters such as recombination that will ultimately influence the rate at which resistance may evolve and spread. Here, we performed a genetic cross between two divergent strains of *H. contortus*, and subsequently used whole-genome resequencing of a female worm and her brood to identify the distribution of genome-wide variation that characterizes these strains. Using a novel bioinformatic approach to identify variants that segregate as expected in a pseudotestcross, we characterized linkage groups and estimated genetic distances between markers to generate a chromosome-scale F<sub>1</sub> genetic map. We exploited this map to reveal the recombination landscape, the first for any helminth species, demonstrating extensive variation in recombination rate within and between chromosomes. Analyses of these data also revealed the extent of polyandry, whereby at least eight males were found to have contributed to the genetic variation of the progeny analyzed. Triploid offspring were also identified, which we hypothesize are the result of nondisjunction during female meiosis or polyspermy. These results expand our knowledge of the genetics of parasitic helminths and the unusual life-history of *H. contortus*, and enhance ongoing efforts to understand the genetic basis of resistance to the drugs used to control these worms and for related species that infect livestock and humans throughout the world. This study also demonstrates the feasibility of using whole-genome resequencing data to directly construct a genetic map in a single generation cross from a noninbred nonmodel organism with a complex lifecycle.

**Key words:** F<sub>1</sub> genetic map, *Haemonchus contortus*, ploidy, polyandry, pseudotestcross, recombination landscape.

## Introduction

Recombination is a key genetic process: the breaking and rejoining of genetic material to produce novel genotypes and in turn, generate phenotypic variation. In eukaryotes, this is achieved by crossing-over between homologous chromosomes during the generation of gametes in meiosis. A common approach to studying recombination is to perform controlled matings (i.e., genetic crosses) between genetically distinct and inbred parents. The parents and offspring are then genotyped to construct genetic linkage maps, which aim to order genes or genetic markers based on the recombination frequency between them. This approach can also be used to identify regions of the genome underlying phenotypic variation, and has been widely used for mapping both simple and complex traits in a range of different organisms (Valentim et al. 2013; Zamanian et al. 2017). More recently, as whole-genome sequencing data have become available for many organisms, genetic maps have been used to inform or validate contig order in genome assemblies (Srinivasan et al. 2002, 2003; Opperman et al. 2008; Nemetschke et al. 2010; Thomas et al. 2012). Where a contiguous genome assembly is already available, a linkage map can be used to explore variation in recombination rates throughout the genome (Rockman and Kruglyak 2009) and determine how this has shaped other aspects of genome architecture, such as the distribution of repeats or the impact of natural selection.

Understanding variation in the rate and pattern of recombination is crucially important in the design and analysis of experiments aimed at mapping the genetic basis of phenotypic traits and in interpreting genetic variation in natural populations. Between species, a negative relationship between genome size and recombination rate has been described (Lynch 2006). Within a species, variation in recombination rate is strongly influenced by the sex of the organism; recombination may not occur in one of the two sexes (typically the heterogametic sex, i.e., the Haldane–Huxley rule; Burt et al. 1991), or, if recombination does occur in both sexes, then females tend to exhibit higher recombination rates than males (i.e., heterochiasmy; Lenormand and Dutheil 2005). In addition, recombination rates have been shown to vary considerably within and between chromosomes, which has been attributed to genomic features including but not limited to GC content, gene density, gene size, simple repeats, and chromatin state (Barnes et al. 1995; Beye et al. 2006; Chan et al. 2012; Tortoreau et al. 2012). Among nematodes, recombination is best characterized in the model organism *Caenorhabditis elegans*, where direct comparison of the physical and genetic maps clearly reveals asymmetrically distributed high and low recombination rate domains in each chromosome, correlated with low and high gene density (and gene expression), respectively (Barnes et al. 1995; Rockman and Kruglyak 2009; Kaur and Rockman 2014). However, even in this model organism, the precise local DNA features that mediate these rate changes remain unclear. Less is

known about recombination in parasitic helminths. Low density genetic maps are available for only three species, the root knot nematode *Meloidogyne hapla* (Opperman et al. 2008; Thomas et al. 2012), the human blood-fluke *Schistosoma mansoni* (Criscione et al. 2009), and the rat gastrointestinal parasite *Strongyloides ratti* (Nemetschke et al. 2010), and only discrete regions of recombination variation have been described in *M. hapla* (Thomas et al. 2012). Recombination rate variation has been proposed to influence the distribution of genetic variation, and in turn, evolution of phenotypic traits in *C. elegans* (Cutter and Payseur 2003; Rockman et al. 2010; Andersen et al. 2012). Therefore, understanding genome-wide recombination variation in parasitic species will likely be important in predicting the genetic architecture and evolution of important parasite life history traits, including pathogenicity, response to host immunity, and chemotherapeutic selection.

The parasite *Haemonchus contortus* is among the most pathogenic of the gastrointestinal nematodes and exerts significant burdens on animal health and the economic viability of livestock farming (Urquhart 1996). It is also an emerging model for the biology of parasitic helminths more widely, particularly for understanding anthelmintic drug action and resistance (Gilleard 2013). *H. contortus* makes a particularly good model because 1) it is the most genetically tractable of any of the strongylid (clade V) parasitic nematodes, a large and important group of parasites including key human and veterinary pathogens; 2) it is a sexually reproducing diploid organism for which the karyotype—five autosomes and XX/XO sex chromosomes—is well defined (Bremner 1954; Redman et al. 2008a); 3) two published draft genome sequences and extensive transcriptomic data are available (Laing et al. 2013, 2016; Schwarz et al. 2013); 4) it is highly fecund, making propagation of crosses easier; 5) it is amenable to cryopreservation of isolates; and 6) it is one of the few parasitic nematode species in which genetic crosses have been successfully established (Le Jambre 1977; Le Jambre et al. 1979, 2000, 2005; Sangster et al. 1998; Hunt et al. 2010; Redman et al. 2012).

Anthelmintic drug failure is an important economic and animal health problem, as anthelmintic resistance is widespread on many farms, and populations and isolates resistant to most major broad-spectrum classes of anthelmintics have been described (Kotze et al. 2014; Mederos et al. 2014; Sales and Love 2016). Accordingly, significant research effort is focused on the development of novel anthelmintics (Kaminsky et al. 2008) or vaccines (Bassetto and Amarante 2015) for parasite control. Although research on *H. contortus* has been instrumental in understanding some of the mechanisms by which resistance arises (Kotze et al. 2014; Britton et al. 2016), the genetic basis of resistance remains largely unresolved and is likely complex. For example, while resistance to benzimidazoles—the class of anthelmintics for which the genetic basis of resistance is best understood—has been linked clearly to mutations at three sites in the isotype-1  $\beta$ -tubulin gene (Kwa et al. 1994; Silvestre and Cabaret 2002; Ghisi et al. 2007), there is evidence that it is a

more complex trait than previously assumed (Zamanian et al. 2017). In contrast, genome-wide studies of ivermectin response—another major anthelmintic—in a number of parasitic helminth species support the hypothesis that this is a quantitative, multigenic trait (Bourguinat et al. 2015; Choi et al. 2017; Doyle et al. 2017). Therefore, establishing the genomic context in which drug resistance alleles are inherited using *H. contortus* will help to resolve the mechanisms by which resistance evolves and spreads in other species of parasitic nematodes as well.

The purpose of this study was to produce a genetic map of *H. contortus*, initially in order to establish an anchored framework for a draft genome under development, and subsequently to estimate the frequency and distribution of recombination in the genome. To do so, we performed a cross between two genetically divergent strains of *H. contortus* that differed in their anthelmintic resistance phenotypes: one that was fully susceptible (Redman et al. 2008b) and one that showed high levels of resistance to three commonly used anthelmintics (Williamson et al. 2011). Four constraints restrict use of *H. contortus* crosses to implement standard classical approaches for genetic mapping: 1) there is an extremely high level of sequence polymorphism present both in field and laboratory strains of *H. contortus* (Gilleard and Redman 2016); 2) few very highly inbred isolates are available to use as parents, and so isolates comprise multiple genotypes; 3) it is difficult, although not impossible, to perform single parent crosses from inbred lines (Sargison et al. 2017); and 4) mating is polyandrous, that is, multiple males can, and will, mate with a single female (Redman et al. 2008a). To overcome these limitations, we developed a genomic strategy to characterize segregating single nucleotide polymorphisms within families based on variants present in a single female and her progeny, which was subsequently used to construct an F<sub>1</sub> genetic map and explore the recombination landscape of the genome. We discuss the implications of recombination, and other novel life history traits identified here, in the context of generating and maintaining genetic variation in parasite populations, and how these factors might impact the spread of anthelmintic resistance in this species.

## Materials and Methods

### Ethics Approval and Consent to Participate

All experimental procedures described in this manuscript were examined and approved by the Moredun Research Institute Experiments and Ethics Committee and were conducted under approved UK Home Office licenses in accordance with the Animals (Scientific Procedures) Act of 1986. The Home Office licence number is PPL 60/03899 and experimental code identifier was E46/11.

### Construction of the Genetic Cross and Collection of Worm Samples

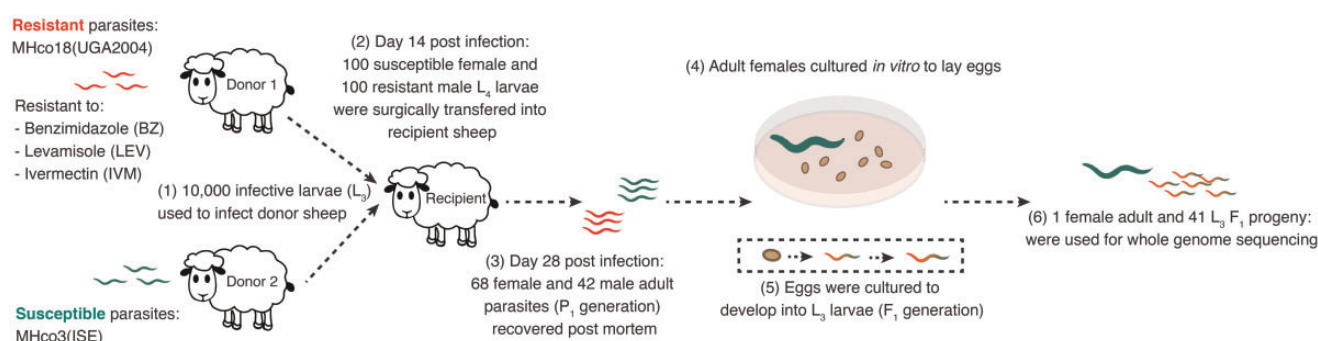
A schematic of the experimental genetic cross is outlined in figure 1. Briefly, two parasite naïve lambs were each infected

with ~10,000 infective larvae from one of two ovine-derived *H. contortus* strains, the anthelmintic susceptible MHco3(ISE) (Redman et al. 2008b), or MHco18(UGA2004) (Williamson et al. 2011), a multidrug resistant strain that is insensitive to standard manufacturers recommended dose rates of benzimidazole, imidazothiazole, and macrocyclic lactone anthelmintics. At 14 days postinfection (DPI), developing sexually immature parasitic stages were recovered post mortem, and the sex of the L<sub>4</sub> stage immature adults was determined by microscopic examination of gross morphology (Denham 1969; Ministry of Agriculture Fisheries and Food 1971). A total of 100 MHco3(ISE) female and 100 MHco18(UGA2004) male L<sub>4</sub> (P<sub>1</sub> generation) were surgically transferred into the abomasum of a recipient sheep to allow reproduction that would generate F<sub>1</sub> hybrid progeny between the two strains. At 28 DPI, 67 MHco3(ISE) females and 42 male MHco18(UGA2004) P<sub>1</sub> from the recipient sheep were recovered post mortem, after which the males were snap frozen in liquid nitrogen and stored. Sampling was performed at 28 DPI to ensure that all of the females would have mated, and that they would be mature enough to have more viable progeny than is thought to be the case in early patency. Individual females were placed into individual wells of 24-well cluster plates (Sarstedt) containing 1 ml of warm RPMI 1640 cell culture media containing 1% (v/v) D-glucose, 2 mM glutamine, 100 IU/ml penicillin, 100 mg/ml streptomycin, 125 mg/ml gentamycin, 25 mg/ml amphotericin B (Redmond et al. 2006), and Hepes (1% v/v) and incubated in 5% CO<sub>2</sub> at 37 °C for 48 h to promote egg shedding. Eggs were transferred at 24 and 48 h and mixed with fresh helminth egg-free sheep faeces before being incubated at 24 °C for 2 weeks to allow larval development to L<sub>3</sub>. After this time, a single female parent (P<sub>1</sub>) and a total of 41 F<sub>1</sub> L<sub>3</sub> progeny were individually stored in preparation for DNA extraction and sequencing library preparation.

### Sample Preparation and Sequencing

The female parent was dissected on ice to isolate the head from the body (in three sections, as three technical replicates) to avoid contamination with fertilized eggs present *in utero*. The female sections and individual L<sub>3</sub> were transferred into 10 µl of sample lysis buffer (working solution: 1000 µl Direct PCR Lysis Reagent [Viagen, Los Angeles], 50 µl 1 M DTT, 10 µl 100 mg/ml Proteinase K) in a 96-well plate and allowed to incubate at 60 °C for 2 h followed by 85 °C for 45 min. Whole genome amplification (WGA) of each sample lysate was performed using RepliG amplification. First, 2–5 µl of sample lysate was combined with 5 µl of 1.3 M Trehalose in a 96-well plate and mixed by gentle tapping, incubated for 3 min at 95 °C, and placed on ice. A 40-µl REPLI-g amplification mix (29 µl REPLI-g Reaction Buffer + 1 µl REPLI-g polymerase + 10 µl 1.3 M Trehalose) was added to each well, and incubated for 16 h at 30 °C followed by 10 min at 65 °C before being placed on ice. The WGA DNA was cleaned using Ampure





**Fig. 1.**—Outline of genetic cross between MHco3(ISE) drug susceptible and MHco18(UGA2004) multidrug resistant *Haemonchus contortus*. A total of 68 MHco3(ISE) females and 42 MHco18(UGA2004) males (from an infection of 100 individuals of each sex) were recovered post mortem ( $P_1$  generation), after which reproductively mature females were incubated in vitro to lay eggs that were subsequently cultured to  $L_3$  stage. These larvae represent the  $F_1$  generation of the cross.

XP beads at a 1.4× bead: DNA reaction ratio, before being eluted in 50  $\mu$ l of RNase/DNase-free water and stored at 4 °C.

PCR-free sequencing libraries (mean length of ~400 bp) were prepared by methods previously described (Kozarewa et al. 2009) and sequenced on an Illumina HiSeq X10, resulting in  $\sim 3.06 \times 10^9$  151-bp paired-end reads (see [supplementary table S1, Supplementary Material](#) online, for a breakdown of reads per lane and per sample). Raw sequence data are archived under the ENA study accession ERP024253.

### Mapping and Variant Analysis

Raw sequence data were mapped to the current unpublished version of the reference genome for *Haemonchus contortus* (v3.0, available at [http://ngs.sanger.ac.uk/production/pathogens/Haemonchus\\_contortus](http://ngs.sanger.ac.uk/production/pathogens/Haemonchus_contortus)) using *Smalt* (<http://www.sanger.ac.uk/science/tools/smalt-0>) with the mapping parameters “-y 0.8 -i 800.” Data from multiple sequencing lanes for a single sample were merged (*samtools-1.3 merge*) and duplicate reads removed (*Picard v2.5.0*; <https://github.com/broadinstitute/picard>) from the bam files before further processing.

Variants were called using *GATK Unified Genotyper* (v3.3.0) (McKenna et al. 2010). The raw variant set was initially filtered to flag variants as low quality if they met the following conditions: quality by depth (QD) < 2; Fisher’s test of strand bias (FS) > 60; RMS mapping quality (MQ) < 40; rank sum of alt versus reference mapping quality (MQRankSum) < -12.5; read position rank sum (ReadPosRankSum) < 8; read depth (DP) < 10. Variants were filtered further using *vcftools* (v0.1.14) (Danecek et al. 2011) to exclude sites with low quality flags, minimize loci with missing data (“—max-missing 0.8”), exclude indels (“—remove-indels”), exclude SNPs with genotype quality (GQ) < 30, and ensure sites were biallelic (“—min-alleles 2, —max-alleles 2”). A gff file generated from *RepeatMasker* of the reference genome was also used to filter variants from the vcf file that were likely associated with repetitive and difficult to map regions of the genome.

Sex determination of the  $F_1$  progeny was performed by measuring: 1) the relative ratio of autosome to X chromosome (characterized and thus named based on synteny with *C. elegans* autosomes and X chromosome) read depth calculated using *samtools-1.3 bedcov*; and 2) the relative heterozygosity of the X chromosome using *vcftools* (v0.1.14) “—het.”

### Genetic Map Construction

A “pseudotestcross” (PT) strategy (Grattapaglia and Sederoff 1994) was employed to generate the genetic map, which required that each input variant site was: 1) heterozygous in the female parent, and 2) segregating in a 1:1 genotype ratio in the  $F_1$  progeny. The segregation pattern of each SNP was first calculated in the  $F_1$  progeny (with “A” referring to the reference allele and “a” to the nonreference variant allele), which resulted in SNPs being placed into one of four categories that best described the likely genotypes of the parents of the cross for that given SNP: 1) “PT: 110,” that is, AA × Aa, 2) “PT: 011,” that is, Aa × aa, 3) “intercross,” that is, Aa × Aa, or 4) SNPs that were clearly segregating in the brood, but for which the segregation ratio of genotypes in the progeny did not fit a simple Mendelian segregation pattern that could be generated via reproduction from a single pair of parents. SNP density was further reduced using *vcftools* (v0.1.14) (Danecek et al. 2011)—*thin* as described in the text. The number of filtered SNPs per segregation group is described in [supplementary table S2, Supplementary Material](#) online. Genotypes for autosomal PT: 011 and PT: 110 SNPs that were heterozygous in the female parent were imported into *R-3.2.2* (R Core Team 2015), after which pairwise recombination fractions (RF) and logarithm of the odds (LOD) scores were determined for each chromosome using *R/QTL* (Broman et al. 2003). Recombination fractions were converted into map distance in centimorgans (cM) using the kosambi map function. Variants resulting in inflation of map distances were identified using *qtl::droponemarker*, and as outliers relative to surrounding markers via visual inspection of LOD and RF using

*qtlncharts::iPlot* (Broman 2015). These aberrant markers were removed in the generation of the final map.

A reverse cross design, whereby SNPs were chosen that: 1) segregated in a 1:1 genotype ratio and 2) were homozygous in the female parent (and therefore putatively heterozygous in the male parents) was also performed. Although polyandry prevented a male-specific genetic map from being constructed (multiple male parents confounded the calculation of linkage between heterozygous sites), these data were used to determine the segregation frequency of alleles from the male parents.

### Recombination Landscape

Recombination patterns for each chromosome were visualized first by generating genotype matrices of pseudotestcross markers for each chromosome using *vcftools* (v0.1.14) “—012,” followed by visualization using the *gplots::heatmap2* function in R. These maps highlighted recombination breakpoints, linkage blocks, and regions of excess heterozygosity or reduced heterozygosity. Recombination rate changes throughout the genome were visualized by constructing Marey maps, which compare the position of the marker in the genome (base position in the fasta sequence) to the relative position in the genetic map. A fitted loess smoothed line of the genetic map positions in 1-Mb windows was performed to calculate the recombination rate.

### Kinship Analysis

Analysis of genetic relatedness between  $F_1$  progeny was undertaken to characterize evidence of polyandry and to determine, if present, the impact on the cross analysis. Principal component analysis (PCA) of genetic distances between the  $F_1$  progeny and female parent was performed using the *SNPrelate* package in R 3.1.2 (Zheng et al. 2012). Kinship coefficients were determined for all pairwise relationships among the  $F_1$  progeny using *KING* (Manichaikul et al. 2010). Relationship networks of the pairwise kinship coefficients were visualized using *Gephi* (v0.9.1; Bastian et al. 2009) to highlight full- and half-sib relationships among the  $F_1$  progeny. Layout of the kinship network graph was determined using the *Force Atlas* parameter, with the nodes ( $F_1$  individuals) colored by their proposed kinship group, and the thickness of the edges proportionate to the kinship coefficient between two  $F_1$  individuals (nodes).

## Results

### Genome Sequencing and Genetic Diversity of a Genetic Cross between Two Isolates of *H. contortus*

A genetic cross was performed between two genetically and phenotypically defined *H. contortus* strains: females were from MHco3(ISE), a serially passaged anthelmintic susceptible “laboratory” strain that has been well characterized by genomic and transcriptomic analyses (Laing et al. 2013, 2016), and

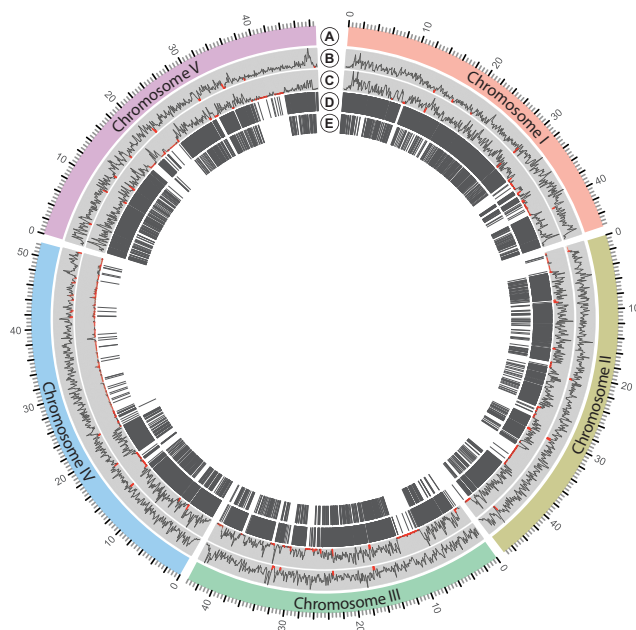
males were from MHco18(UGA2004), a multidrug resistant serially passaged strain originally isolated from the field at the University of Georgia (Williamson et al. 2011) (fig. 1). Whole genome sequencing (WGS) was performed on DNA derived from a single adult MHco3(ISE) female parent and 41 of her  $F_1$   $L_3$  progeny to achieve a minimum 30 $\times$  sequencing coverage per sample (mean sequencing depth: 34.80 $\times$   $\pm$  16.16 SDs), generating a median yield of 65.97 million reads per sample (supplementary table S1, Supplementary Material online). Mapping of the sequencing data was performed using an improved genome assembly of the MHco3(ISE) isolate described by Laing et al. (2013), which now consists of five scaffolds representing the autosomal chromosomes and two scaffolds representing the X chromosome, for an assembly length of  $\sim$ 279 Mb. Sequence depth of the X chromosome scaffolds relative to the five autosomal scaffolds, together with rates of heterozygosity on the X chromosome scaffolds, revealed 20 male and 21 female  $F_1$  progeny in the brood.

Approximately 5.3 million single nucleotide polymorphisms (SNPs) that passed stringent filtering criteria were identified in the autosomal chromosomes (fig. 2A and supplementary table S2, Supplementary Material online), at a genome-wide density of 2,242 SNPs per 100 kb (fig. 2B), or 1 SNP per 44.6 base pairs (bp). A pseudotestcross approach was used to generate the  $F_1$  genetic map, which required that candidate markers: 1) were heterozygous in the female parent; and 2) segregated in a ratio statistically indistinguishable from a 1:1 genotype ratio in the  $F_1$  progeny. By using these criteria, we identified a set of markers that could be analyzed using the same statistical approaches as conventional linkage mapping using a test cross. Analysis of the 730,825 heterozygous SNPs in the female MHco3(ISE) parent demonstrated that the distribution of variation was not uniform throughout the genome, with a number of long contiguous regions of homozygosity observed (fig. 2C and supplementary fig. S1, Supplementary Material online). In particular,  $\sim$ 27 Mb of the second half of chromosome IV was largely homozygous, containing  $\sim$ 50% more homozygous variant sites and  $\sim$ 30% less heterozygous sites compared with the genome-wide average (supplementary table S3, Supplementary Material online).

Among the SNPs that were heterozygous in the female parent, 171,876 SNPs segregated at an  $\sim$ 1:1 genotype ratio in the  $F_1$  progeny (supplementary table S2, Supplementary Material online; PT: 110 and PT: 011). To avoid including tightly linked SNPs, the 171,876 candidate SNPs were thinned to 1 per 25,000 bp, which resulted in a final candidate list of 5,595 SNPs for analysis in the cross.

### Characterization of an Autosomal $F_1$ Genetic Map Generated Using Pseudotestcross SNP Markers

Initial analysis of genome-averaged genotype ratios (supplementary fig. S2, Supplementary Material online) of the candidate pseudotestcross sites in each  $F_1$  individual revealed that



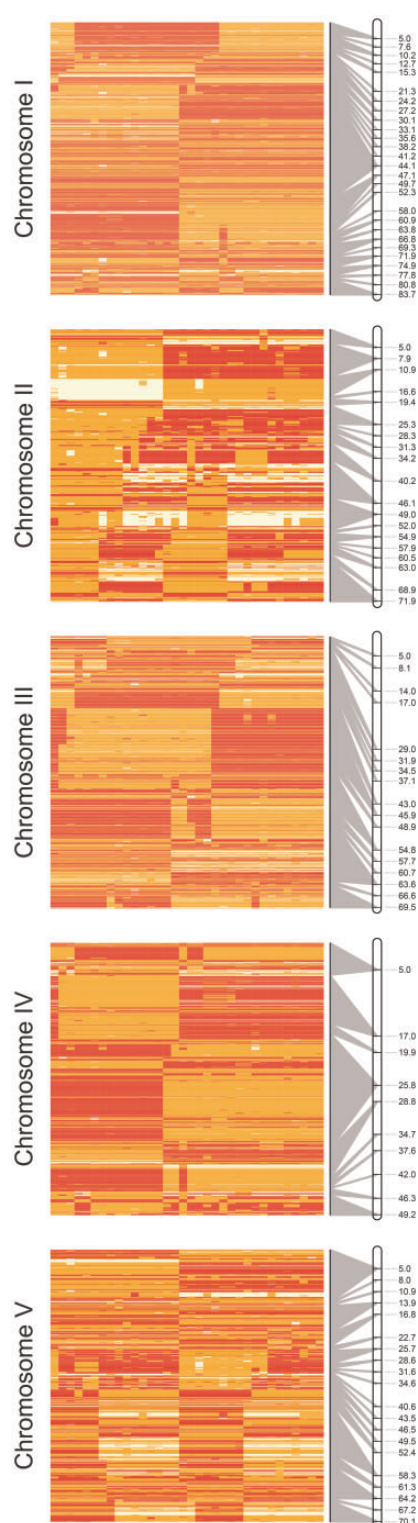
**FIG. 2.**—Autosome-wide variant density and candidate genetic map markers identified from the female parent and  $F_1$  progeny. (A) The five autosomes of *Haemonchus contortus*, named based on synteny with *Caenorhabditis elegans* chromosomes, span 237 Mb. (B) SNP density was calculated in 100-kb windows, and is presented as the relative variant density of the female parent ( $P_1$ ) and all  $F_1$  progeny. (C) Density of heterozygous variants in the female parent. (D) Positions of candidate pseudotestcross SNPs that were heterozygous in the female parent and segregated in a 1:1 genotype ratio in the  $F_1$  progeny. Red annotations in plots (C) and (D) highlight low density regions, defined as genome-wide mean SNP density minus 3 SD. (E) Positions of the final set of 1,618 SNPs used in the  $F_1$  genetic map. The plot was produced using Circos (Krzywinski et al. 2009).

most individuals displayed an approximate 50:50 ratio of homozygous: heterozygous genotypes, as expected. However, seven individuals presented as outliers with an excess of heterozygous genotypes (supplementary fig. S2A, Supplementary Material online; moderate outliers: individuals F1\_12, F1\_30, F1\_40; extreme outliers: individuals F1\_21, F1\_23, F1\_32, F1\_38). The variant-allele frequency distribution of these individuals (supplementary fig. S3, Supplementary Material online) revealed a skew consistent with a nondiploid complement of chromosomes, with a major nonreference (relative to the genome assembly) allele frequency peak at  $\sim 30\%$  and minor peak at 60% frequency. This allele frequency skew was typically found across all chromosomes within an individual, suggesting that they were not aneuploids. A notable exception was individual F1\_30 (one of the moderate outliers), where chromosomes I, III, and V had a distinct allele frequency spectrum consistent with more than two copies of each chromosome present, relative to chromosomes II and IV, which appeared to be disomic. All seven of these nondiploid individuals were therefore removed from the

pseudotestcross analysis (supplementary fig. S2B and D, Supplementary Material online;  $n = 34$ ).

A reanalysis of the remaining 34 individuals revealed 217,575 pseudotestcross SNPs, 129,985 intercross SNPs, and 383,265 SNPs that were heterozygous in the female parent but did not segregate in a way compatible with analysis as a single-pair mating cross (supplementary table S2, Supplementary Material online). Thus, a total of 4,587 pseudotestcross SNPs (217,575 SNPs thinned to 1 SNP per 25,000 bp) were candidate markers for the map construction using R/QTL (fig. 2D), from which 1,618 SNPs were used in the final genetic map (fig. 2E and supplementary table S4, Supplementary Material online). Recombination plots and genetic maps for the five autosomes are presented in figure 3, and characteristics of the map are presented in table 1. The total map distance of the five autosomes was  $\sim 344.46$  cM. The number of markers per chromosome ranged from 215 on chromosome II to 475 on chromosome I, with a mean value of 323.6 markers per chromosome. Significant gaps in the map correlated with absence, or very low density, of the prerequisite heterozygous SNPs in the female parent, as described earlier (fig. 2C). This loss of markers was most obvious in chromosome IV, where only approximately half of the chromosome is represented in the map, resulting in a map length of 49.21 cM, compared with the average map length of other chromosomes of 73.79 cM. The genome-wide recombination rate was on an average 604.12 ( $\pm 84.01$  SD) kb/cM or 1.68 ( $\pm 0.25$  SD) cM/Mb, which corresponded to an overall average number of crossover events per chromosome of 0.69 ( $\pm 0.12$  SD). Chromosome IV was again an outlier, with a recombination rate of 2.01 cM/Mb,  $\sim 21\%$  higher than the other four autosomes (1.68 cM/Mb average). Analysis of the X chromosome diversity from the adult female and all progeny revealed 100,016 SNPs in the 23.3- and 18.9-Mb X-linked scaffolds; this frequency (1 SNP per 422 bp) equates to  $\sim 10$ -fold fewer variable sites on the X chromosome relative to the autosomes. Attempts to generate an X chromosome genetic map were limited by a lack of prerequisite heterozygous variant sites in the female X chromosome sequences (supplementary fig. S1, Supplementary Material online). To explore this further, the diversity of hemizygous genotypes called in the male  $F_1$  progeny, that is, genotyped as AA or aa reflecting the haploid  $X^A$  or  $X^a$  allele, respectively, was compared with genotypes resolved in the female parent (supplementary fig. S4, Supplementary Material online). Strikingly, male genotypes were entirely concordant with the female parent, further supporting the lack of segregating genetic diversity in the female parent diploid X chromosomes. Female  $F_1$  progeny contained both homozygous and heterozygous sites in their X chromosomes; given the lack of variation in the female parent, this diversity was entirely inherited from the paternal X chromosome.





**FIG. 3.**—Recombination and genetic maps of the five autosomes of *Haemonchus contortus*. Recombination plots depict genotype segregation patterns per  $F_1$  progeny (columns; clustered by genetic similarity) of pseudotestcross markers used in the genetic map (rows). Segregating “parental” and “recombinant” haplotypes inherited from the female parent are indicated by opposing color schemes. Genotypes: AA: red;

### Patterns of Recombination within Autosomal Chromosomes of the $F_1$ Progeny

Analysis of recombination rate throughout each chromosome was determined by comparing physical and genetic distances, which can be visualized in a Marey map (Chakravarti 1991) (fig. 4). Recombination rate (fig. 4 red line; cM/Mb) was not uniform throughout the chromosomes, nor was it consistent between chromosomes. Chromosomes I, II, and IV tended to show a pattern of three main recombination rate domains; a reduced recombination rate domain toward the middle of the chromosome, flanked by domains of increased recombination rate that extend toward the ends of the chromosomes. This three-domain pattern was not as clear for chromosomes III and V; chromosome III showed a greater recombination rate in the first half of the chromosome that decreased throughout the second half of the chromosome, whereas chromosome V had longer low recombination rate domains toward the ends of the chromosome arms, and greater recombination rate toward the middle of the chromosome. It is curious that chromosome IV retained the three-domain recombination architecture, given that the right arm is largely missing due to lack of the prerequisite heterozygous sites in this region of the female parent (fig. 2C and [supplementary fig. S1, Supplementary Material](#) online). Each chromosome also showed evidence of additional low recombination rate domains at one or both ends of the chromosome in the subtelomeric regions extending into the chromosome. Finally, within the elevated recombination rate domains, the recombination rate was not necessarily constant; discrete peaks of high recombination rates were observed in all chromosomes. However, the relative positions of these high recombination peaks were not the same between chromosomes.

### Family Structure and Kinship among the Brood

*Haemonchus contortus* is known to be polyandrous (Redman et al. 2008a). This knowledge, together with the observation that >50% of SNPs did not segregate in either a 1:1 or 1:2:1 genotype ratio ([supplementary table S2, Supplementary Material](#) online), suggested that the 41 progeny analyzed were sired from more than a single male parent. An initial analysis of genetic relatedness by principal component analysis (PCA) of 21,822 autosomal SNPs (complete data set thinned using a linkage disequilibrium threshold of 0.5 and minor allele frequency of 0.05) revealed obvious genetic structure, with at least four (PC 1 vs. 2) to as many as six (PC 2 vs. 3)

### FIG. 3. Continued

Aa: yellow; aa: white. The relationship between SNP position in the recombination map and genetic map position (cM) is represented by a connecting gray line; multiple SNPs between which no recombination was observed collapse into a single map position in the genetic map (gray ribbon from multiple SNPs to a single map marker).



**Table 1**

Summary Characteristics of the F<sub>1</sub> Genetic Map, Including Number of Markers Used, Map Length, Recombination Rate, and Crossover Frequency

Chromosome	Chromosome Length (bp)	Markers Used (#)	Genetic Map Length (cM)	Recombination Rate (kb/cM) <sup>a</sup>	Recombination Rate (cM/Mb) <sup>b</sup>	Crossovers per Chromosome <sup>c</sup>
I	45778363	475	83.71	546.87	1.83	0.84
II	47384193	215	71.88	660.13	1.51	0.72
III	43564237	363	69.53	626.55	1.60	0.70
IV <sup>d</sup>	51819793	226	49.21	490.85	2.04	0.49 <sup>e</sup>
V	48825595	339	70.13	696.22	1.44	0.70
Total/average	237372181	1618	344.46	604.12	1.68	0.69

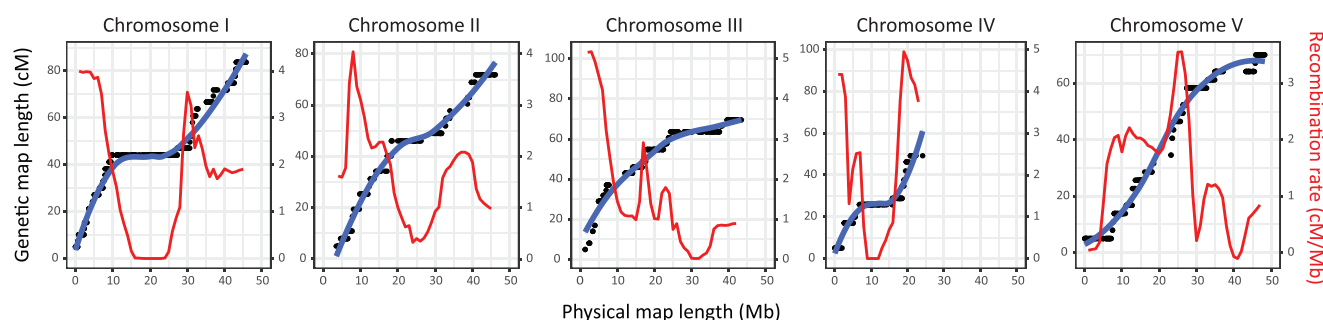
<sup>a</sup>Recombination rate (kb/cM): chromosome length (kb)/genetic map length.

<sup>b</sup>Recombination rate (cM/Mb): genetic map length/(chromosomal length/10<sup>6</sup>).

<sup>c</sup>Crossovers per chromosome: (genetic map length/100)/number of chromosomes.

<sup>d</sup>The genetic map only spanned ~24 Mb of chromosome IV due to homozygosity in the female parent. As such, recombination rates have been calculated for chromosome IV using 24154752 bp (position of the genetic map marker closest to the homozygosity region) as the chromosome length.

<sup>e</sup>Likely to be underestimated given only half of the chromosome is present.



**FIG. 4.**—Analysis of recombination rate variation throughout the genome. Marey maps were constructed to show the relationship between the genetic position of each marker (black point) relative to the physical position of the marker in the genome. Line of best fit was plotted using default parameters of the *geom\_smooth* function of *ggplot2* in R. Recombination rates (cM/Mb; red line) were calculated by calculating genetic map distance in 1-Mb windows throughout the genome from a fitted *loess*-smoothed line of the genetic map positions.

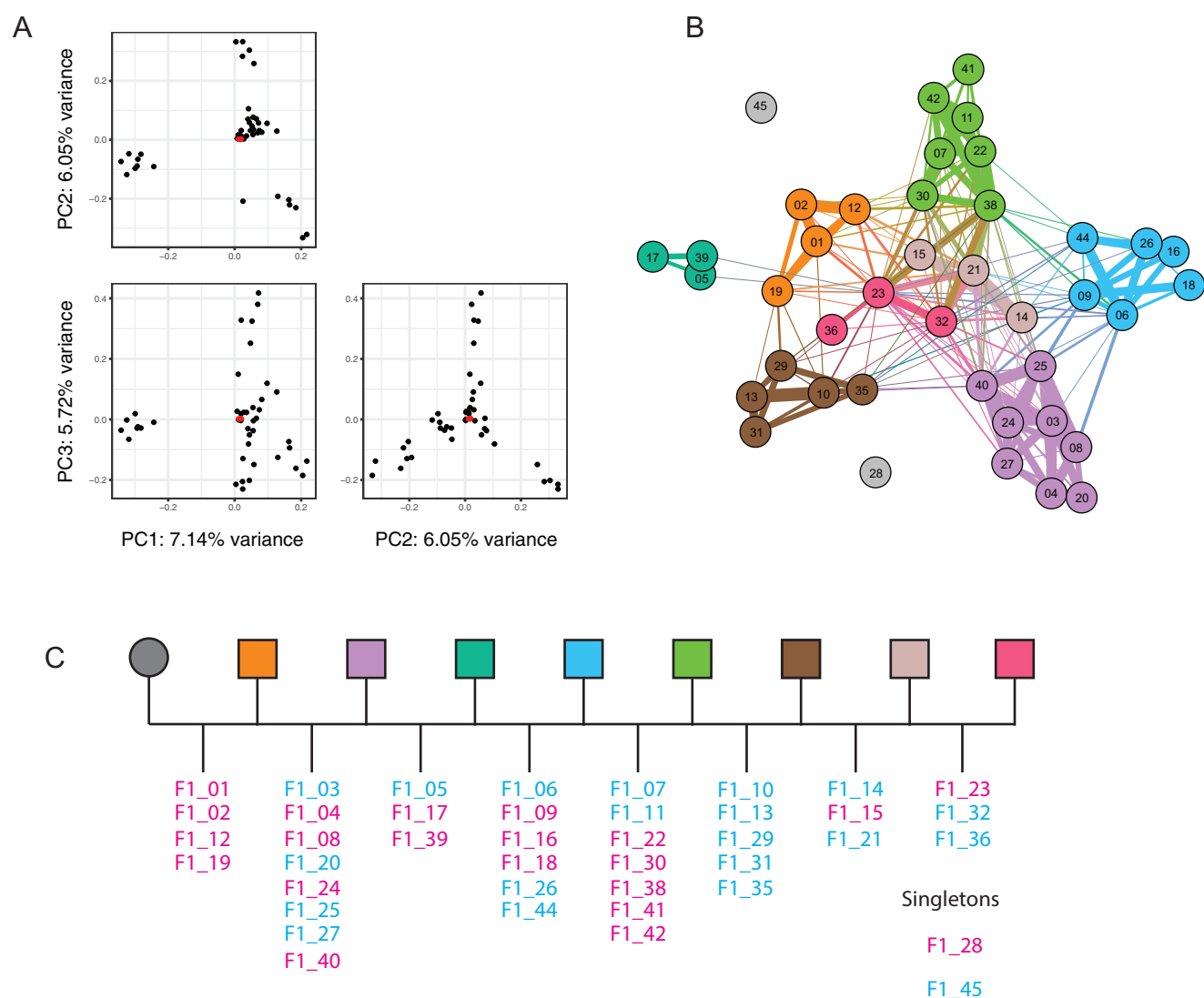
putative clusters of F<sub>1</sub> progeny (fig. 5A), consistent with the hypothesis that the brood resulted from polyandrous mating.

To more accurately describe these putative relationships among the progeny, we calculated kinship coefficients (Manichaikul et al. 2010), which describe the probability that a given allele in two individuals is identical by descent (i.e., an allele shared due to recent shared ancestry, as opposed to identical by state, in which the allele is simply shared by two individuals without common ancestry), for all pairwise combinations of progeny. Employing all autosomal SNPs ( $n = 5,323,039$  SNPs), this analysis revealed eight clusters of full-sib relationships containing multiple F<sub>1</sub> progeny (fig. 5B). Two individuals, F1\_28 and F1\_45, did not share any pairwise kinship coefficients consistent with a full-sib relationship with any individual, and hence, may represent the progeny from additional paternal contributions to the brood. Three individuals, F1\_21, F1\_23, and F1\_38, seemed to show full-sib relationships with individuals from multiple families via strong kinship associations between themselves and others. Intriguingly, these were the same individuals identified as outliers with excess heterozygosity (supplementary fig. S2, Supplementary Material online) and that showed a skewed

allele frequency distribution (supplementary fig. S3, Supplementary Material online) suggestive of aneuploidy or polyploidy. These autosomal kinship data are further supported by the observation that X chromosome diversity in the female F<sub>1</sub> progeny, which reflects paternal X chromosome inheritance in the absence of maternal X chromosome diversity, clusters the female F<sub>1</sub> progeny into five groups of two or more individuals (supplementary fig. S4, Supplementary Material online). Three unclustered individuals were also identified for the X chromosomes, including individual F1\_28, which did not share any full-sib relationships in the kinship analysis (fig. 5B). These X chromosome derived clusters are concordant with the full-sib family structure using autosomal SNPs. Taken together, these data describing the familial relationships among the F<sub>1</sub> progeny cohort lead us to propose a pedigree consisting of at least eight paternal contributions (fig. 5C).

## Discussion

Our comprehensive genetic characterization of genome-wide patterns of segregation in progeny from a brood of parasites



**Fig. 5.**—Familial relationships determined via analysis of genetic diversity and kinship between full- and half-sibs. (A) Principal component analysis of parent and progeny genetic diversity, comparing the top three principal components (PCA). The female parental values ( $n = 3$ ) are indicated as red points in each plot. (B) Network analysis of kinship coefficients determined by *KING* (Manichaikul et al. 2010) and visualized by *Gephi* (Bastian et al. 2009) highlighting full-sib relationships between progeny. The thickness of the line (edges) represents the kinship coefficient between individuals (nodes) and is proportionate to the relationship between pairs. (C) Proposed pedigree of the brood. Full-sib male (blue) and female (pink) progeny are indicated for each subfamily. Colors used in (B) and (C) represent groups of progeny that share a common father.

revealed extensive variation in recombination rates across chromosomes, and confirmed previous suggestions of polyandry as the dominant mating system in *H. contortus* (Redman et al. 2008a). Moreover, analysis of genetic variation in both autosomes and the X chromosome identified extended regions of reduced heterozygosity in the female parent, which could be a genetic consequence of population bottlenecks during the generation and maintenance of the MHco3(ISE) line. Analysis of allele frequency spectra also suggested the presence of polyploids among the progeny. The availability of a largely complete chromosomal scale *H. contortus* genome assembly facilitated such analyses. Here, we

discuss some of the characteristics and challenges associated with the assembly of a genetic map when homozygous single parent crosses are not available, and how some of the features of the genetic cross impact on our understanding of *H. contortus* biology and anthelmintic resistance.

### Prediction of Genomic Structure

A small number of genetic linkage maps have been described for free-living nematodes (*Caenorhabditis elegans* and *Pristionchus pacificus*), parasitic nematodes (*Meloidogyne hapla*, *Strongyloides ratti*), or parasitic trematodes

(*Schistosoma mansoni*). *Haemonchus contortus* was found to have the lowest genome-wide recombination rate among these worms, at an average of 1.68 cM/Mb throughout the ~279-Mb genome. However, the recombination rate tended to scale proportionately with genome size, that is, larger genomes have lower recombination rates (supplementary fig. S5, Supplementary Material online). Although the recombination rates of these worms are somewhat lower than predicted by a model describing the relationship between eukaryotic genome size and recombination rate (supplementary fig. S5, Supplementary Material online, gray dashed line) (Lynch 2006), they are more consistent with recombination rates of other invertebrates (supplementary fig. S5, Supplementary Material online, gray points; see Supplementary Table 1 from Lynch 2006 for invertebrate recombination rate data). This relationship between genome size and recombination rate is somewhat dependent on the number of crossovers per chromosome per meiosis; for example, in *C. elegans*, almost complete crossover interference occurs, such that only a single crossover per pair of homologous chromosome is observed (Meneely et al. 2002). In *H. contortus*, some but certainly not complete interference was observed, with an average rate of 0.69 crossovers per chromosome (i.e., 1.38 crossovers per pair of homologous chromosomes). Given that the genetic map of *P. pacificus* (a free-living clade V nematode often compared with *C. elegans* but more distantly related to *C. elegans* than *H. contortus*) is expanded relative to *C. elegans* due to the presence of double crossovers (Srinivasan et al. 2002; Hong and Sommer 2006), the observation of complete chromosome interference in *C. elegans* is likely to be an exception rather than the rule. The mechanisms by which recombination rates in free living or parasitic species are controlled is largely unknown. However, it is clear that there is significant diversity among the helminths: from single to few crossovers per homologous chromosomes per meiosis in *C. elegans* and *H. contortus*/*P. pacificus*, respectively, to multiple chiasmata between homologous pairs in *S. mansoni* (Hirai et al. 1996), and in an extreme case, recombination between all four chromatids within a homologous pair in *M. hapla* (Liu et al. 2007). The presence of such diversity provides an insight into the evolutionary potential of these species, and perhaps an opportunity to explore the mechanisms of recombination rate variation via comparative genomic approaches.

To our knowledge, we are the first to report the use of whole genome sequencing to construct a genetic map of any helminth species. WGS allowed significantly greater flexibility in choosing high quality variants to be included in the genetic map than other marker-based approaches such as amplified fragment length or Sanger-sequencing derived markers, and more recently, higher throughput RADseq and genotype-by-sequencing approaches, and allowed us to fully exploit the genetic variation in the available progeny. This was particularly important given that: 1) the progeny were not derived from a

cross between genetically distinct homozygous single male and female parents, as is typical for a genetic mapping experiment; 2) the high genetic diversity within isolates meant that a lot of markers have to be screened and discarded to find “biallelic markers” that segregate appropriately for analysis; and 3) we did not know how many males would contribute to the progeny of the cross due to polyandry. As such, we developed a bioinformatics pipeline to select markers based on the genotype segregation ratio of the progeny (approximate 1:1 genotype ratios: Aa: aa [PT: 011] or AA: Aa [PT: 110]) and heterozygous sites in the female. This unusual cross design, chosen to account for the biological complexity, meant that relatively few of the sites that differed between parents were usable in the map (pseudotestcross SNPs represent only 4.09% of the total SNPs in the brood, and 29.77% of SNPs heterozygous in the female parent, before deliberate thinning). A very large panel of traditional markers would thus have been required even for the relatively small number of progeny analyzed here. The genome-wide resequencing approach that we used would seem to be the only practical way to generate complete recombination maps in this system. Genome-wide genetic variation that has been validated as segregating in a Mendelian fashion also provides a valuable resource for downstream experiments such as: QTL analyses of parasite traits (e.g., drug resistance); using individuals phenotyped in vitro using bioassays (Le Jambre 1976; Coles et al. 1988; Hunt and Taylor 1989; Álvarez-Sánchez et al. 2005); or as a source of genome-wide population genetic markers, which typically require low/no linkage disequilibrium between loci.

We initially intended to use the F<sub>1</sub> genetic map to guide improvements of the assembly of the draft genome for *H. contortus* MHco3(ISE) (Laing et al. 2013); while subsequent improvements to the genome assembly have rendered this unnecessary, the colinearity of the genetic and physical maps confirms the accuracy of the current assembly. A number of features of this data set would not have been obvious without integrating the genetic map and physical assembly. The first of these includes the nonuniform distribution of genetic map markers in the genome. This is most obvious in chromosome IV in which approximately half of the chromosome is missing from the genetic map, due to a long tract of homozygosity in the female parent. However, each chromosome contained multiple megabase-scale gaps that directly corresponded to a deficiency of heterozygosity in the female parent in these regions. This may reflect the genetic history of this particular strain: MHco3(ISE) is a laboratory strain (Redman et al. 2008b) derived from ISE, a strain that was originally generated by 15 rounds of half-sib matings of an outbred isolate (Roos et al. 2004). Since that time, MHco3(ISE) has been passaged and cryopreserved on numerous occasions at an unknown, but likely limited, population size. Although significant diversity remains in this strain (Redman et al. 2008b; Sargison et al. 2017), it is probable that population bottlenecks, increased

inbreeding, or selection have resulted in discrete regions of the genome becoming genetically fixed. Secondly, the integration of the genetic map and contiguous physical genome map allowed us to describe the recombination landscape of the genome. Although there are similarities in the recombination rate domain structure with that of *C. elegans* (Barnes et al. 1995; Rockman and Kruglyak 2009), chromosomes III and V have distinct recombination rate differences compared both to chromosomes I, II, and IV of *H. contortus*, and to all chromosomes of *C. elegans*. The broad-scale distribution is unlikely to be the result of differential recombination around centromeric sequences, given the similarities in recombination domain structure with *C. elegans* chromosomes, and that *C. elegans* chromosomes are holocentric during mitosis (Albertson and Thomson 1993; Wicky and Rose 1996). However, it has been proposed that the low or absent recombination in the chromosome termini may correlate with the presence of a spindle attachment site that guides segregation of homologous chromosomes in meiosis (Cutter et al. 2009). Although we have no data to directly test whether *H. contortus* is holocentric, we have identified low recombining chromosome termini consistent with that observed in *C. elegans*.

Despite the relatively high marker density used here ( $n = 1,618$ ), many SNPs were completely linked in seemingly nonrecombining regions. Inclusion of a larger number of progeny would provide additional resolution to more precisely characterize variation in and transitions between recombination rate domains in each chromosome. Finally, although we could not generate a genetic map for the X chromosome due to the limited brood size and the absence of genetic diversity in the female parent, WGS data allowed us to examine genetic diversity among the female progeny, which highlighted both significant genetic variation and clustering consistent with shared paternal haplotypes in the autosomes.

### Detection of Polyandry

Technical challenges associated with single male and female mating led us to perform the genetic cross using 100 immature female MHco3(ISE) and 100 male MHco18(UGA2004) surgically implanted into the abomasum of a recipient sheep. Analysis of the genetic diversity among  $F_1$  progeny of a single female revealed discrete groups of progeny; given that *H. contortus* has been previously described to be polyandrous (Redman et al. 2008a), we hypothesized that these groups represented the progeny of different male nematodes. In this cross, our data support at least eight paternal genotypes contributing to multiple individuals in the brood ( $n = 41$ ). These data are consistent with the original report of polyandry in *H. contortus*, which described at least three to four paternal microsatellite-derived genotypes from the 11 to 17 progeny sampled per single fecund female analyzed (Redman et al. 2008a). Single worm genotyping of males recovered from the initial genetic cross recipient lamb would provide further

insight into the ancestral relationships among the progeny. The remarkably high frequency of polyandrous pairings would substantially increase the diversity of genotypes found among the progeny, as more possible pairs of haplotypes would be generated. This feature of *H. contortus* biology is likely to play a significant role in generating and maintaining the high levels of genetic diversity characterized in laboratory (Redman et al. 2008b) and field (Silvestre et al. 2009; Redman et al. 2015) isolates of this parasite and is also relevant to other parasitic nematode species where polyandry has been reported (Grillo 2005; Zhou et al. 2011; Hildebrandt et al. 2012).

### Detection of Nondiploid Patterns of Variation

*Haemonchus contortus* is a dioecious, sexually reproducing diploid animal. Unexpectedly, we observed 7 of the 41 progeny (17.1%) with an excess of heterozygous genotypes, and with an allele frequency spectrum that is consistent with a polyploid complement of chromosomes. Moreover, two distinct patterns of allele frequency spectrum among six of the seven putative polyploids lead us to hypothesize that these progeny arose by either: 1) nondisjunction during meiosis 1 of gametogenesis in the female parent; or 2) polyspermy, that is, an egg that has been fertilized by more than one sperm, as a consequence of polyandry (see [supplementary fig. S6](#) and [table S5, Supplementary Material](#) online, for alternate hypotheses and evidence for the generation of triploid progeny in the brood). A third hypothesis—nondisjunction during male gametogenesis resulting in diploid sperm—was excluded; analysis of genotype frequencies among the  $F_1$  progeny at SNPs at which the female parent was homozygous demonstrated that paternally derived alleles from putatively heterozygous sites were segregating independently, resulting in an approximate 1:1 genotype ratio among all but one individual ([supplementary fig. S2C, Supplementary Material](#) online; the putative aneuploid  $F1_{30}$ ). This supports the observation that polyploidy was inherited from diploid gametes derived from the female parent (i.e., nondisjunction), or multiple haploid gametes from the male parents (i.e., polyspermy).

Polyploidy has been previously described among nematodes. In *C. elegans*, a range of ploidy states have been characterized (see Hodgkin 2005 for review of work on natural and induced tetraploids, triploids, and haploids) and is a feature of a cellular organismal growth into late adulthood due to nuclei endoreduplication (Hedgecock and White 1985; Flemming et al. 2000). However, polyploidy is typically associated with parthenogenesis in worms (e.g., some *Meloidogyne* spp.; Lunt et al. 2014; Blanc-Mathieu et al. 2017; and some *Panagrolaimus* spp. Schiffer et al. 2017). Polyspermy in worms is thought to be rare, with a single description in the rodent filarial worm *Acanthocheilonema viteae* (McLaren 1973); more is understood in regard to the mechanisms by which polyspermy is prevented (Jaffe 1976; Wong and Wessel 2005; Johnston et al. 2010). However,



polyspermy may be associated with polyandrous mating (Arnqvist and Rowe 2005), whereby sexual conflict among males (at least eight in the data presented) competing to reproduce with a female likely results in strong selection on male reproductive traits (e.g., sperm count, size, and quality), which increases the likelihood of reproductive success (Birkhead and Pizzari 2002). Although this would drive coevolution of female traits to block polyspermy, it may be that polyspermy is a consequence of this competition in polyandrous species such as *H. contortus*. Given that these progeny were sampled at the L<sub>3</sub> stage, we cannot be sure that these individuals would have developed to adulthood and become reproductively viable. However, a report describing the karyotype of a single triploid *H. contortus* adult female suggests that they may be at least developmentally viable (Bremner 1954). The presence of sporadic polyploidy among the *H. contortus* F<sub>1</sub> progeny represents a novel finding among parasitic nematodes; further work is required to determine if triploidy is a feature of *H. contortus* biology and prevalent in the field, or, is a novel feature of this genetic cross. If the former is true, then it will be important to be aware of ploidy variation in population genetic studies of *H. contortus*, particularly if larval stages are sampled.

A single individual (F1\_30) presented with a variant allele frequency spectrum consistent with an aneuploid complement of chromosomes. Aneuploidy and other severe chromosomal abnormalities have been described in experimental hybrid crosses between *H. contortus* and the related cattle parasite, *Haemonchus placei* (Le Jambre and Royal 1980); such hybrids have recently been genetically characterized in the field (Chaudhry et al. 2015). Although such chromosomal abnormalities have not been described in within-species *H. contortus* crosses to date, the use of whole genome sequencing provides greater resolution over single marker techniques to detect these chromosome-wide changes, which may have resulted via incompatibility of rare alleles between the genetically diverse strains used in the cross.

## Conclusions

In summary, we have undertaken a comprehensive analysis of genetic diversity within a *H. contortus* family derived from an experimental genetic cross. Whole-genome sequencing of a female and her brood allowed the construction of a F<sub>1</sub> genetic map, despite the challenging design dictated by the unusual biology and life history of this parasitic helminth. Development of the genetic map continues to build upon the genetic resources available for *H. contortus* as an experimentally tractable parasitic species, and provides new insight into the recombination architecture of the genome. Moreover, the close relationship between *H. contortus* and *C. elegans* argues for further comparative studies between these species. These data, together with evidence of polyandry and polyploidy, highlight the complexities of the

underlying biology of *H. contortus*, and have important implications toward understanding the evolution of anthelmintic resistance in this important pathogen of livestock. Clear recombination rate differences throughout the genome will influence the rate by which a locus correlated (i.e., a genetic marker linked to resistance), or causally associated (i.e., resistance conferring mutation) with anthelmintic resistance will evolve within a population, dependent on the position in the genome that the given locus lies. Incorporating recombination rate parameters in studies that aim to genetically detect or track the transmission of resistance will be critical to the utility and interpretation of data derived from such approaches. This will be particularly the case given the likely multigenic nature of resistance to some, and perhaps all, anthelmintics.

## Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

## Acknowledgments

We acknowledge funding from the Biotechnology and Biological Sciences Research Council (BBSRC) (BB/M003949), Wellcome Trust through their core support of the Wellcome Trust Sanger Institute (206194), and the Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS) for supporting work carried out at Moredun Research Institute. We would like to acknowledge the BUG Consortium members and Parasite Genomics group (WTSI) for helpful comments and suggestions toward this work, Pathogen Informatics and DNA Pipelines (WTSI) for their support and expertise, Taisei Kikuchi for the whole-genome amplification protocol for nematode larvae, Ray Kaplan for supplying L<sub>3</sub> larvae from the MHco18(UGA2004) strain, and to the Bioservices Division, Moredun Research Institute, for expert care and assistance with animals.

## Literature Cited

- Albertson DG, Thomson JN. 1993. Segregation of holocentric chromosomes at meiosis in the nematode, *Caenorhabditis elegans*. *Chromosome Res.* 1(1):15–26.
- Álvarez-Sánchez MA, Pérez García J, Bartley D, Jackson F, Rojo-Vázquez FA. 2005. The larval feeding inhibition assay for the diagnosis of nematode anthelmintic resistance. *Exp Parasitol.* 110(1):56–61.
- Andersen EC, et al. 2012. Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat Genet.* 44(3):285–290.
- Arnqvist G, Rowe L. 2005. *Sexual conflict*. Princeton (NJ): Princeton University Press.
- Barnes TM, Kohara Y, Coulson A, Hekimi S. 1995. Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*. *Genetics* 141(1):159–179.
- Bassetto CC, Amarante AF. 2015. Vaccination of sheep and cattle against haemonchosis. *J Helminthol.* 89(5):517–525.
- Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. *ICWSM* 8:361–362.

- Beye M, et al. 2006. Exceptionally high levels of recombination across the honey bee genome. *Genome Res.* 16(11):1339–1344.
- Birkhead TR, Pizzari T. 2002. Postcopulatory sexual selection. *Nat Rev Genet.* 3(4):262–273.
- Blanc-Mathieu R, et al. 2017. Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. *PLoS Genet.* 13(6):e1006777.
- Bourguinat C, et al. 2015. Macrocyclic lactone resistance in *Dirofilaria immitis*: failure of heartworm preventives and investigation of genetic markers for resistance. *Vet Parasitol.* 210(3–4):167–178.
- Bremner KC. 1954. Cytological polymorphism in the nematode *Haemonchus contortus* (Rudolphi 1803) Cobb 1898. *Nature* 174(4432):704–705.
- Britton C, Roberts B, Marks ND. 2016. Functional genomics tools for *Haemonchus contortus* and lessons from other helminths. *Adv Parasitol.* 93:599–623.
- Broman KW. 2015. R/qtlcharts: interactive graphics for quantitative trait locus mapping. *Genetics* 199(2):359–361.
- Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19(7):889–890.
- Burt A, Bell G, Harvey PH. 1991. Sex differences in recombination. *J Evol Biol.* 4(2):259–277.
- Chakravarti A. 1991. A graphical representation of genetic and physical maps: the Marey map. *Genomics* 11(1):219–222.
- Chan AH, Jenkins PA, Song YS, McVean G. 2012. Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. *PLoS Genet.* 8(12):e1003090.
- Chaudhry U, et al. 2015. Genetic evidence for hybridisation between *Haemonchus contortus* and *Haemonchus placei* in natural field populations and its implications for interspecies transmission of anthelmintic resistance. *Int J Parasitol.* 45(2–3):149–159.
- Choi YJ, et al. 2017. Genomic introgression mapping of field-derived multiple-anthelmintic resistance in the nematode parasite *Teladorsagia circumcincta*. *PLoS Genet.* 13(6):e1006857.
- Coles GC, Tritschler JP II, Giordano DJ, Laste NJ, Schmidt AL. 1988. Larval development test for detection of anthelmintic resistant nematodes. *Res Vet Sci.* 45:50–53.
- Criscione CD, Valentim CL, Hirai H, LoVerde PT, Anderson TJ. 2009. Genomic linkage map of the human blood fluke *Schistosoma mansoni*. *Genome Biol.* 10(6):R71.
- Cutter AD, Dey A, Murray RL. 2009. Evolution of the *Caenorhabditis elegans* genome. *Mol Biol Evol.* 26(6):1199–1234.
- Cutter AD, Payseur BA. 2003. Selection at linked sites in the partial selfer *Caenorhabditis elegans*. *Mol Biol Evol.* 20(5):665–673.
- Danecek P, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158.
- Denham DA. 1969. The development of *Ostertagia circumcincta* in lambs. *J Helminthol.* 43(3):299–310.
- Doyle SR, et al. 2017. Genome-wide analysis of ivermectin response by *Onchocerca volvulus* reveals that genetic drift and soft selective sweeps contribute to loss of drug sensitivity. *PLoS Negl Trop Dis.* 11(7):e0005816.
- Flemming AJ, Shen ZZ, Cunha A, Emmons SW, Leroi AM. 2000. Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. *Proc Natl Acad Sci U S A.* 97(10):5285–5290.
- Ghisi M, Kaminsky R, Maser P. 2007. Phenotyping and genotyping of *Haemonchus contortus* isolates reveals a new putative candidate mutation for benzimidazole resistance in nematodes. *Vet Parasitol.* 144(3–4):313–320.
- Gilleard JS. 2013. *Haemonchus contortus* as a paradigm and model to study anthelmintic drug resistance. *Parasitology* 140(12):1506–1522.
- Gilleard JS, Redman E. 2016. Genetic Diversity and Population Structure of *Haemonchus contortus*. *Adv Parasitol.* 93:31–68.
- Grattapaglia D, Sederoff R. 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* 137(4):1121–1137.
- Grillo V. 2005. Development of microsatellites and the population genetic analysis of the parasitic nematode *Teladorsagia circumcincta*. [PhD Thesis] University of Glasgow, UK.
- Hedgecock EM, White JG. 1985. Polyploid tissues in the nematode *Caenorhabditis elegans*. *Dev Biol.* 107(1):128–133.
- Hildebrandt JC, Eisenbarth A, Renz A, Streit A. 2012. Single worm genotyping demonstrates that *Onchocerca ochengi* females simultaneously produce progeny sired by different males. *Parasitol Res.* 111(5):2217–2221.
- Hirai H, Hirata M, Aoki Y, Tanaka M, Imai HT. 1996. Chiasma analyses of the parasite flukes, *Schistosoma* and *Paragonimus* (Trematoda), by using the chiasma distribution graph. *Genes Genet Syst.* 71(3):181–188.
- Hodgkin J. 2005. Karyotype, ploidy and gene dosage, *WormBook*, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.3.1, <http://www.wormbook.org>.
- Hong RL, Sommer RJ. 2006. *Pristionchus pacificus*: a well-rounded nematode. *Bioessays* 28(6):651–659.
- Hunt KR, Taylor MA. 1989. Use of the egg hatch assay on sheep faecal samples for the detection of benzimidazole resistant nematodes. *Vet Rec.* 125(7):153–154.
- Hunt PW, et al. 2010. The use of DNA markers to map anthelmintic resistance loci in an intraspecific cross of *Haemonchus contortus*. *Parasitology* 137(04):705–717.
- Jaffe LA. 1976. Fast block to polyspermy in sea-urchin eggs is electrically mediated. *Nature* 261(5555):68–71.
- Johnston WL, Krizus A, Dennis JW. 2010. Eggshell chitin and chitin-interacting proteins prevent polyspermy in *C. elegans*. *Curr Biol.* 20(21):1932–1937.
- Kaminsky R, et al. 2008. A new class of anthelmintics effective against drug-resistant nematodes. *Nature* 452(7184):176–180.
- Kaur T, Rockman MV. 2014. Crossover heterogeneity in the absence of hotspots in *Caenorhabditis elegans*. *Genetics* 196(1):137–148.
- Kotze AC, et al. 2014. Recent advances in candidate-gene and whole-genome approaches to the discovery of anthelmintic resistance markers and the description of drug/receptor interactions. *Int J Parasitol Drugs Drug Resist.* 4(3):164–184.
- Kozarewa I, et al. 2009. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat Methods* 6(4):291–295.
- Krzywinski MI, et al. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res.* 19(9):1639–1645.
- Kwa MS, Veenstra JG, Roos MH. 1994. Benzimidazole resistance in *Haemonchus contortus* is correlated with a conserved mutation at amino acid 200 in beta-tubulin isotype 1. *Mol Biochem Parasitol.* 63(2):299–303.
- Laing R, et al. 2013. The genome and transcriptome of *Haemonchus contortus*, a key model parasite for drug and vaccine discovery. *Genome Biol.* 14(8):R88.
- Laing R, et al. 2016. *Haemonchus contortus*: genome structure, organization and comparative genomics. *Adv Parasitol.* 93:569–598.
- Le Jambre LF. 1976. Egg hatch as an in vitro assay of thiabendazole resistance in nematodes. *Vet Parasitol.* 2(4):385–391.
- Le Jambre LF. 1977. Genetics of vulvar morph types in *Haemonchus contortus*: *Haemonchus contortus cayugensis* from the Finger Lakes Region of New York. *Int J Parasitol.* 7(1):9–14.
- Le Jambre LF, Geoghegan J, Lyndal-Murphy M. 2005. Characterization of moxidectin resistant *Trichostrongylus colubriformis* and *Haemonchus contortus*. *Vet Parasitol.* 128(1–2):83–90.
- Le Jambre LF, Gill JH, Lenane IJ, Baker P. 2000. Inheritance of avermectin resistance in *Haemonchus contortus*. *Int J Parasitol.* 30(1):105–111.

- Le Jambre LF, Royal WM. 1980. Meiotic abnormalities in backcross lines of hybrid *Haemonchus*. *Int J Parasitol.* 10(4):281–286.
- Le Jambre LF, Royal WM, Martin PJ. 1979. The inheritance of thiabendazole resistance in *Haemonchus contortus*. *Parasitology* 78(2):107–119.
- Lenormand T, Dutheil J. 2005. Recombination difference between sexes: a role for haploid selection. *PLoS Biol.* 3(3):e63.
- Liu QLL, Thomas VP, Williamson VM. 2007. Meiotic parthenogenesis in a root-knot nematode results in rapid genomic homozygosity. *Genetics* 176(3):1483–1490.
- Lunt DH, Kumar S, Koutsovolous G, Blaxter ML. 2014. The complex hybrid origins of the root knot nematodes revealed through comparative genomics. *PeerJ* 2:e356.
- Lynch M. 2006. The origins of eukaryotic gene structure. *Mol Biol Evol.* 23(2):450–468.
- Manichaikul A, et al. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics* 26(22):2867–2873.
- McKenna A, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20(9):1297–1303.
- McLaren DJ. 1973. Oogenesis and fertilization in *Dipetalonema viteae* (Nematoda: Filarioidea). *Parasitology* 66(3):465–472.
- Mederos AE, Ramos Z, Banchero GE. 2014. First report of monepantel *Haemonchus contortus* resistance on sheep farms in Uruguay. *Parasit Vectors* 7:598.
- Meneely PM, Farago AF, Kauffman TM. 2002. Crossover distribution and high interference for both the X chromosome and an autosome during oogenesis and spermatogenesis in *Caenorhabditis elegans*. *Genetics* 162(3):1169–1177.
- Ministry of Agriculture Fisheries and Food. 1971. Manual of veterinary parasitological laboratory techniques. London: H.M. Stationery Off. p. vi, 131 p.
- Nemetschke L, Eberhardt AG, Viney ME, Streit A. 2010. A genetic map of the animal-parasitic nematode *Strongyloides ratti*. *Mol Biochem Parasitol.* 169(2):124–127.
- Opperman CH, et al. 2008. Sequence and genetic map of *Meloidogyne hapla*: a compact nematode genome for plant parasitism. *Proc Natl Acad Sci U S A.* 105(39):14802–14807.
- R Core Team. 2015. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.
- Redman E, et al. 2008a. Genetics of mating and sex determination in the parasitic nematode *Haemonchus contortus*. *Genetics* 180(4):1877–1887.
- Redman E, et al. 2008b. Microsatellite analysis reveals marked genetic differentiation between *Haemonchus contortus* laboratory isolates and provides a rapid system of genetic fingerprinting. *Int J Parasitol.* 38(1):111–122.
- Redman E, et al. 2012. Introgression of ivermectin resistance genes into a susceptible *Haemonchus contortus* strain by multiple backcrossing. *PLoS Pathog.* 8(2):e1002534.
- Redman E, et al. 2015. The emergence of resistance to the benzimidazole anthelmintics in parasitic nematodes of livestock is characterised by multiple independent hard and soft selective sweeps. *PLoS Negl Trop Dis.* 9(2):e0003494.
- Redmond DL, et al. 2006. An immunogenic cathepsin F secreted by the parasitic stages of *Teladorsagia circumcincta*. *Int J Parasitol.* 36(3):277–286.
- Rockman MV, Kruglyak L. 2009. Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet.* 5(3):e1000419.
- Rockman MV, Skrovanek SS, Kruglyak L. 2010. Selection at linked sites shapes heritable phenotypic variation in *C. elegans*. *Science* 330(6002):372–376.
- Roos MH, Otsen M, Hoekstra R, Veenstra JG, Lenstra JA. 2004. Genetic analysis of inbreeding of two strains of the parasitic nematode *Haemonchus contortus*. *Int J Parasitol.* 34(1):109–115.
- Sales N, Love S. 2016. Resistance of *Haemonchus* sp. to monepantel and reduced efficacy of a derquantel/abamectin combination confirmed in sheep in NSW, Australia. *Vet Parasitol.* 228:193–196.
- Sangster NC, Redwin JM, Bjorn H. 1998. Inheritance of levamisole and benzimidazole resistance in an isolate of *Haemonchus contortus*. *Int J Parasitol.* 28(3):503–510.
- Sargison ND, et al. 2017. A method for single pair mating in an obligate parasitic nematode. *Int J Parasitol.* doi: 10.1016/j.ijpara.2017.08.010.
- Schiffer PH, et al. 2017. Signatures of the evolution of parthenogenesis and cryptobiosis in the genomes of panagrolaimid nematodes. *bioRxiv.* doi: 10.1101/159152.
- Schwarz EM, et al. 2013. The genome and developmental transcriptome of the strongylid nematode *Haemonchus contortus*. *Genome Biol.* 14(8):R89.
- Silvestre A, Cabaret J. 2002. Mutation in position 167 of isotype 1 beta-tubulin gene of Trichostrongylid nematodes: role in benzimidazole resistance? *Mol Biochem Parasitol.* 120(2):297–300.
- Silvestre A, Sauve C, Cortet J, Cabaret J. 2009. Contrasting genetic structures of two parasitic nematodes, determined on the basis of neutral microsatellite markers and selected anthelmintic resistance markers. *Mol Ecol.* 18(24):5086–5100.
- Srinivasan J, et al. 2002. A bacterial artificial chromosome-based genetic linkage map of the nematode *Pristionchus pacificus*. *Genetics* 162(1):129–134.
- Srinivasan J, et al. 2003. An integrated physical and genetic map of the nematode *Pristionchus pacificus*. *Mol Genet Genomics* 269(5):715–722.
- Thomas VP, et al. 2012. A sequence-anchored linkage map of the plant-parasitic nematode *Meloidogyne hapla* reveals exceptionally high genome-wide recombination. *G3 (Bethesda)* 2(7):815–824.
- Tortoreau F, et al. 2012. A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. *BMC Genomics* 13(1):586.
- Urquhart GM. 1996. Veterinary parasitology. Oxford (UK); Ames (IA): Blackwell Science.
- Valentim CL, et al. 2013. Genetic and molecular basis of drug resistance and species-specific drug action in schistosome parasites. *Science* 342(6164):1385–1389.
- Wicky C, Rose AM. 1996. The role of chromosome ends during meiosis in *Caenorhabditis elegans*. *Bioessays* 18(6):447–452.
- Williamson SM, et al. 2011. Candidate anthelmintic resistance-associated gene expression and sequence polymorphisms in a triple-resistant field isolate of *Haemonchus contortus*. *Mol Biochem Parasitol.* 180(2):99–105.
- Wong JL, Wessel GM. 2005. Defending the zygote: search for the ancestral animal block to polyspermy. *Curr Top Dev Biol.* 72:1–151.
- Zamanian M, et al. 2017. Discovery of unique loci that underlie nematode responses to benzimidazoles. *bioRxiv.* doi: 10.1101/116970.
- Zheng X, et al. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28(24):3326–3328.
- Zhou CH, Yuan K, Tang XL, Hu NY, Peng WD. 2011. Molecular genetic evidence for polyandry in *Ascaris suum*. *Parasitol Res.* 108(3):703–708.

Associate editor: Michael Lynch